RAG\_Legal\_Docs

Lead-Scoring-Case-Study-Upgrad The main objective of this assignment is to process and analyse a collection text files containing legal agreements (e.g., NDAs) to prepare them for implementing a Retrieval-Augmented Generation (RAG) system. This involves:

1.Understand the Cleaned Data : Gain a comprehensive understanding of the structure, content, and context of the cleaned dataset.

2.Perform Exploratory Analysis : Conduct bivariate and multivariate analyses to uncover relationships and trends within the cleaned data.

3.Create Visualisations : Develop meaningful visualisations to support the analysis and make findings interpretable.

4.Derive Insights and Conclusions : Extract valuable insights from the cleaned data and provide clear, actionable conclusions.

5.Document the Process : Provide a detailed description of the data, its attributes, and the steps taken during the analysis for reproducibility and clarity.

The ultimate goal is to transform the raw text data into a clean, structured, and analysable format that can be effectively used to build and train a RAG system for tasks like information retrieval, question-answering, and knowledge extraction related to legal agreements.